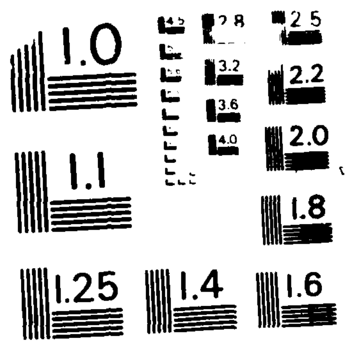
AD-A190 387 NON-RIGID MOTION AND REGGE CALCULUS(U) MASSACHUSETTS INST OF TECH CAMBRIDGE ARTIFICIAL INTELLIGENCE LAB R JASINSCHI ET AL NOV 87 AI-M-996 N00014-85-K-0124 1/1

UNCLASSIFIED F/G 12/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo 996. November, 1987

Non-rigid motion and Regge calculus.

R. Jasinschi and A. Yuille

Abstract. ~~We study~~ the problem of recovering the structure from motion of figures which are allowed to perform a controlled non-rigid motion. We use Regge Calculus to approximate a general surface by a net of triangles. The non-rigid flexing motion ~~we~~ deal with corresponds to keeping the triangles rigid and allowing bending only at the joins between triangles. Such motion has been studied by Koenderink and van Doorn (1986). ~~We show~~ that this motion keeps the Gaussian curvature of the surface constant but changes the principal curvatures.

~~We~~ show that depth information of the vertices of the triangles can be obtained by using a modified version of the Incremental Rigidity Scheme devised by Ullman (1984). In cases where the motion of the figure displays fundamentally different views at each frame presentation the algorithm works well, not only for strictly rigid motion (Ullman 1984, Grzwacz and Hildreth 1985) but also for a limited amount of bending deformation. ~~We modify~~ this scheme, to allow for flexing motion (in the sense defined above) ~~and call~~ our ~~version~~ the Incremental Semirigidity Scheme.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-85-K-0124.

/

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AI Memo 996	2. GOVT ACCESSION NO. ADA190387	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Non-rigid Motion and Regge calculus		5. TYPE OF REPORT & PERIOD COVERED memorandum
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) R. Jasinschi and Alan Yuille		8. CONTRACT OR GRANT NUMBER(s) N00014- 85-K-0124
9. PERFORMING ORGANIZATION NAME AND ADDRESS Artificial Intelligence Laboratory 545 Technology Square Cambridge, MA 02139		10. PROGRAM ELEMENT PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency 1400 Wilson Blvd. Arlington, VA 22209		12. REPORT DATE November 1987
		13. NUMBER OF PAGES 34
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Office of Naval Research Information Systems Arlington, VA 22217		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

structure from motion
incremental rigidity
regge calculus

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We study the problem of recovering the structure from motion of figures which are allowed to perform a controlled non-rigid motion. We use Regge Calculus to approximate a general surface by a net of triangles. The non-rigid flexing motion we deal with corresponds to keeping the triangles rigid and allowing bending only at the joins between triangles. Such motion has been studied by Koenderink and van Doorn (1986). We show that this motion keeps the Gaussian curvature of the surface constant but changes the principal curvatures. (cont. on back)

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block 20 cont.

We show that depth information of the vertices of the triangles can be obtained by using a modified version of the Incremental Rigidity Scheme devised by Ullman (1984). In cases where the motion of the figure displays fundamentally different views at each frame presentation the algorithm works well, not only for strictly rigid motion (Ullman 1984, Grzwacz and Hildreth 1985) but also for a limited amount of bending deformation. We modify this scheme to allow for flexing motion (in the sense defined above) and call our version the Incremental Semirigidity Scheme.

Introduction

In order to make the study of visual perception and information processing more systematic (and also easier), we divide the multitude of visual information faculties into modules which are treated as being more or less independent. Examples of these are stereo, motion, color and shape from shading. This strategy has led to some very fruitful results (Marr 1982).

All modules act on a basic represention of the image consisting of primitive elements which could be perceptually salient features of the image, such as points or lines, or even the intensity values themselves. Low-level vision consists of applying the different modules to these primitive representations to build up a description of the world.

The basic input to a vision system is the intensity changes occuring on a two-dimensional image surface. The visual processing system has to recover the complete three-dimensional description of objects in space, from this primitive set of data. Mathematically, the visual data is given in terms of variables defined on a two-dimensional manifold so that it lacks the necessary information to reconstruct the surfaces of objects embedded in the three-dimensional world. This fundamental issue is given a mathematical basis through the use of the regularization theory (Poggio and Torre 1984). As a consequence of this, additional information has to be furnished to the visual system, usu-

A-1

ally in the form of constraints (which describe basic assumptions about the three-dimensional objects). Examples of these constraints are the rigidity of three-dimensional objects (in visual motion) or surface smoothness (in surface interpolation).

Visual motion, as demonstrated by numerous psychophysical experiments (for reviews see Braddick 1980, Hildreth and Koch 1986), is given in two modes: one is short range, so that information is collected through the use of frames very close in terms of temporal and spatial displacement and the other is long range. In the present paper we will deal mainly with the second type of motion.

Depth from motion can be recovered by seeing the object from different viewpoints. Psychophysical experiments Wallach and O'Connell 1953, Johansson 1975, Wertheimer 1912 have explicitly shown that a subject is able to perceive the complete three-dimensional form of a moving object when presented with different views of it (for both continuous and discrete presentations).

The process of finding the structure of an object from motion information (and its depth) can be divided into a three step process:(i) determining the primitives for the two-dimensional description, (ii) making the correspondence between these primitives and (iii) integrating information between the frames

to get the structure.

The lack of depth information (due to the projection of the object onto the image plane) can be counterbalanced by assuming some sort of constraint for the object. In the case of rigid objects (Ullman 1979) it can be shown that three different views of four non-collinear (arbitrarily chosen) points on the surface of the object gives enough information to recover the motion of the object and hence its depth values.

The first stage of the motion module is to determine the correspondence between features in different frames. These features may include points, line segments or aggregates of them. In this paper we assume that we know the correspondence between features and can track them as they move (in the image plane) through the sucessive (discrete) frames. We must now make assumptions about the object in order to recover its depth. A standard, though strong assumption, is that the object moves rigidly in space (Ullman 1979). We would like to weaken this assumption to allow for more general motion. The Incremental Rigidity Scheme (IRS) (Ullman 1984) is able to recover the structure of a rigidly moving object by assuming the minimal change in rigidity of this object between frames. The IRS is also able to deal with a limited amount of non-rigidity of the object. In this paper we show how the IRS can be (in its modified version) extended to deal with objects undergoing non-

rigid flexing motion preserving the Gaussian curvature. Koenderink and van Doorn (1986) have studied motions of this type. For clarity we refer to the modified IRS as the ISRS (Incremental Semi-rigidity Scheme).

The non-rigid flexing motion we consider corresponds to rigid triangles bending relative to each other. We will show in the next section that this motion corresponds to non-rigid motion which preserves the Gaussian curvature. For example, motion of this type would allow a sheet of paper to be deformed into a cylinder.

The basic idea of the IRS is to contruct a internal model of the object, which is initially choosen to be flat, and to update this model assuming minimal change of rigidity between consecutive image frames. This change in rigidity is measured by the changes in distance between different points of the object. Each new frame yields more information about the object and the scheme converges to a fixed model. For rigid motion this gives good results (Ullman 1984). Our modification of the algorithm requires minimal change of rigidity only for points which lie at adjoining vertices of the triangulation. This is a weaker assumption than global rigidity and, as we shall show, in some cases allows the recovery of structure of objects undergoing highly non-rigid motion.

We work specifically with two types of figures built up of triangles, al-

lowing bending deformations to take place. The first figure is made out of two triangles with a common edge which constitutes the axis of bending. To simulate the non-rigid motion we use a two step proceedure. First we rotate the whole figure as a rigid object and then we bend one triangle with respect to the other (over their common edge). The second figure consists of six triangles with adjacent common edges. Its non-rigid motion, modulo global rigid rotation, is similar to the folding (and unfolding) of a umbrella. Note that for both these examples the triangles are not deformed so the Gaussian curvature remains unchanged.

As an intermediate step we applied the ISRS to the rigid motion of a single triangle and of six triangles with adjacent (common) edges. In both cases, as expected, the algorithm works well. Finally, we studied the six triangle figure including a small deformation of one of the base edges, keeping the others fixed, which corresponds to a deformation which changes the global curvature.

This article is organized as follows:in chapter 2 we give a general overview of the triangulation method, which is known to physicists as "Regge calculus". In chapter 3 the Incremental Rigidity Scheme, and the ISRS, is presented and its application for particular types of triangle aggregates is given in chapter 4. In chapter 5 we discuss some informal psychophysical results we obtained.

Finally in chapter 6 we draw conclusions and indicate future research.

Regge calculus

The basic idea of the Regge calculus (Regge 1961) is to approximate a general surface by a polyhedron built up of triangles. We recover the structure of the general surface as we send the number of triangles to infinity, while maintaining the total area of the polyhedron fixed.

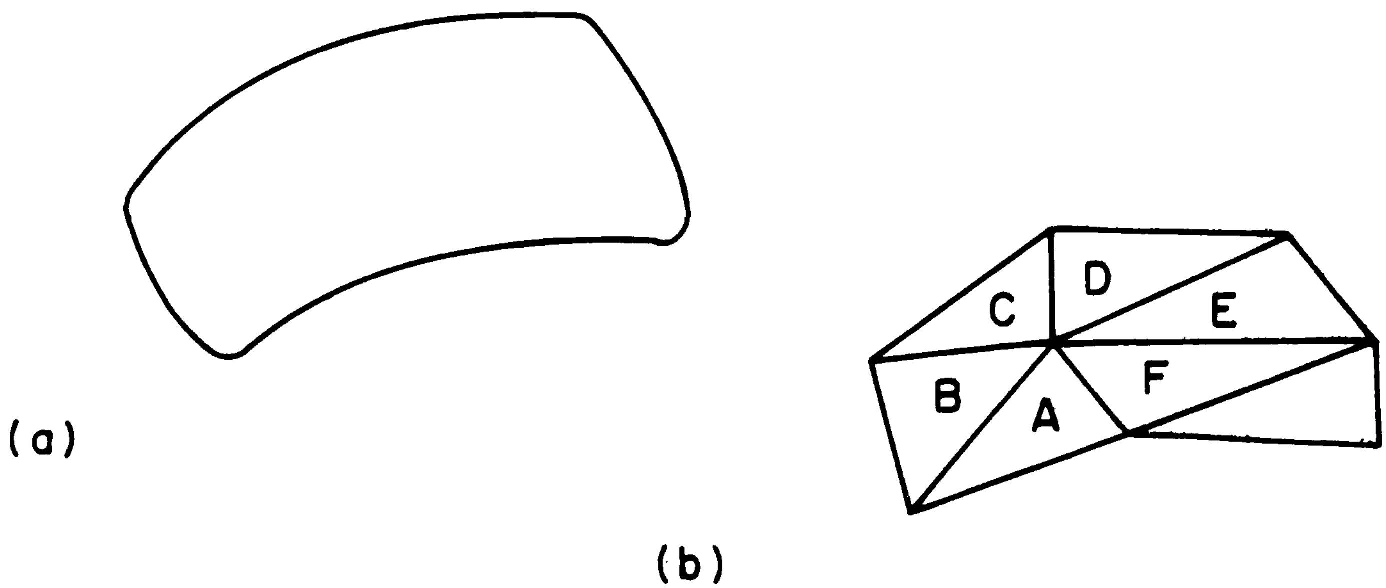


Figure 1 *A general triangulation*

Suppose that we construct a curved triangle on the surface of a sphere, where the edges are geodesics (a geodesic is the line of shortest length between two points). The sum of the internal angles of this triangle is different from π, because the surface is curved and its curvature is given by the difference with

respect to π of the sum of the internal angles. By using a collection of these curved triangles we can reconstruct the surface of the sphere. This means that, at any point of the sphere, the gaussian curvature (and the principal curvatures), using curved triangles, is always the same as the curvature of the sphere.

On the other hand, the Regge calculus builds up a net of flat triangles which only approximate the shape of the initial surface to a given precision (which in turn depends on the scale of triangulation). The triangles used in the triangulation method of the Regge calculus have straight edges so that the curvature content lies exclusively at the vertices (the intersection points of edges of different triangles). Let now us suppose that we approximate the surface of the sphere by a net of triangles (with straight edges). Also, let us concentrate on an arbitrary vertex (intersection of edges) of this triangulation. If we take the difference between 2π and the sum of the angles (adjacent to this vertex) we get the **deficit angle** which measures the local curvature (at the location of the vertex) of the (triangulated) surface.

Let us take a particular vertex α so that the angles of its adjacent triangles (denoted by τ) are represented by Θ_α^τ. The deficit angle δ_α is defined by the following expression

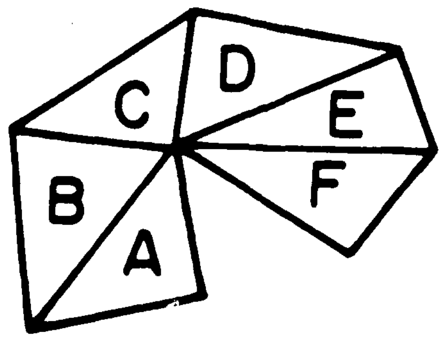


Figure 2 *The deficit angle*

$$\delta_\alpha = 2\pi - \sum_r \Theta_\alpha^r$$

The total curvature of this triangulated surface is given by the sum of all deficit angles. This means, if R is the total curvature, then

$$R = \sum_\alpha \delta_\alpha$$

As a consequence of this, the curvature of a general surface, which must be calculated locally, can be approximated by the sum of the deficit angles for an arbitrary triangulation of it.

In general, the curvature of an arbitrary surface is given by the Euler number χ which describes the topological content of the surface, and is given by

$$\chi = 2 - \eta$$

where η is the surface genus. Let us, as a illustration, think of the surface of a sphere which is approximated by a net of triangles. By using the Euler formula we can write the surface genus as

$$\eta = 2 + e - v - f$$

where e, v and f are respectively the number of edges, vertices and faces of the triangulation net. In order to create a hole we have to eliminate a entire face from this surface, and as a consequence, an equal number of (three) edges and vertices. By simple inspection of equation (4) one can conclude that by creating a hole on the (triangulated) surface the surface genus is reduced by one unit, and, as a consequence of formula (3), the Euler number is reduced by an equal amount. If, on the other hand, we want to create a handle out of the original surface, we have to eliminate two faces and identify the perimeters (constructed from the edges bordering the holes). This means, by analogy to the creation of a hole, that the surface genus (and also the Euler number) is

reduced by two unities. For the general case, if we create a number of B holes and H handles then the surface genus is given by $\eta = B + 2H$, and the Euler number by

$$\chi = 1 - B - 2H$$

The relationship between the curvature (R) and the surface genus η is given by the Gauss-Bonnet curvature theorem (for general compact polyhedrons). It can be writen as

$$R = \sum_{i=1}^{N} \delta_i = 2\pi(1 - \eta)$$

So, for example, when we want to know the total curvature of a sphere, which can be done by summing over all the deficit angles, we just have to observe that its genus is zero (it has no holes or handles) and as result of this we obtain 4π. In the case of a torus, whose genus is 2 (one handle) the total curvature is zero.

We can describe the Regge calculus by the following block diagram which describes a simple algorithm to calculate the curvature of a triangulated surface.

What happens if the size of all triangles in the triangulation net goes to zero at the same rate as their numbers increase? We recover the original

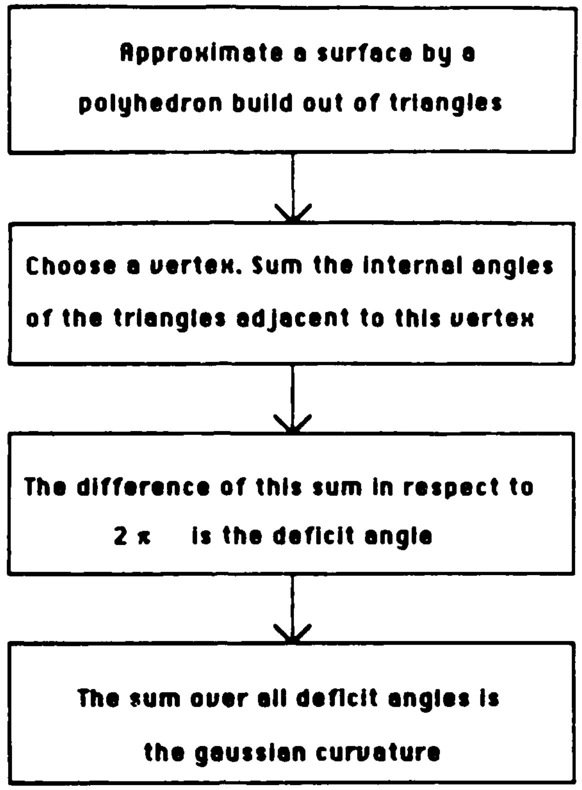


Figure 3 *Regge calculus block diagram*

surface whose gaussian curvature is given by

$$R = \int d^2x \, \mathcal{R}$$

where $\mathcal{R}$ is the local (gaussian) curvature.

The idea of the Regge calculus can be generalized to a n dimensional manifold where it approximates a smoothly curved n-dimensional Riemannian manifold by a collection of n-dimensional elements without any curvature contend (like the triangles in 2-dimensions) joined by (n - 2)-dimensional elements (points in 2-dimensions) where the curvature content is concentrated.

It is certainly easier to calculate the curvature using a triangulation net,

since we simply have to sum over all deficit angles, rather than taking the continuum limit, in which case we have to calculate the curvature of the surface at each point (which is very sensitive to noise).

Our basic idea is to represent objects by the position of the set of points corresponding to a triangulation of the surface. We can track the positions of this set of points through sucessive time frames. It will then be possible to determine the curvature changes in time very simply. More importantly this representation enables us to relax the rigidity assumption and allow for a class of non-rigid motion. Suppose the triangles are kept fixed in size but are allowed to flex relative to each other. Koenderink and Van Doorn (1986) show that although this kind of motion is non-rigid it nonetheless is sufficiently constrained to allow the structure to be recovered. We show that it is straightforward to adapt the incremental rigidity scheme to deal with this type of motion. Thus our method can be thought of as a type of incremental semi-rigidity scheme.

This semi-rigid motion can be easily analysed in terms of Regge Calculus. Since the triangles are fixed, the deficit angles at the vertices are constant. Thus the Gaussian curvature does not change during the motion. Recall that the Gaussian curvature is the product of the two principal curvatures of the object. So, for example, this motion will allow a cylinder to be transformed

into a plane (both have zero Gaussian curvature), but not into a sphere.

Incremental Rigidity

We assume that we have determined the correspondence between the vertices of a triangulation of an arbitrary object and want to determine its structure. Thus we assume the correspondence problem and the triangulation problem are solved and concentrate directly upon determining structure. It could, however, be possible to solve the correspondence problem at the same time as the structure is obtained.

Human visual motion is measured in (at least) two modes, one the **short-range** mode deals with space-time information processed discretely but with a high sampling rate while the other one, the **long-range** mode exibits a low sampling rate. The **long-range** mode has a ISI (interstimulus interval-the temporal interval between two samplings) of at least 300ms in contrast to the **short-range** mode with a ISI less than 80-100ms (Braddick 1980, for a discussion see Marr 1982). Although these two modes act independently in the initial stages of visual motion processing (Braddick 1980), it is supposed that at later stages they have some kind of interaction (Clatworthy and Frisby 1973).

We will assume the analog of the **long-term** mode, which leads to the

perception of apparent motion. The scheme that we adopt for determining structure is based on the Incremental Rigidity Scheme (Ullman 1984) and consists of updating an internal model of the structure of the object. More precisely, we initially assume the object is flat and update the model between time frames by assuming minimal change in rigidity (or semi-rigidity) between frames. This minimal change is enforced by minimizing a cost function, thus yielding a series of estimated depth values, which gradually converge to the correct result.

The IRS can be described as follows. Suppose there are P points. Initially the model assumes that the depth values are zero for all points. Let us suppose that for the Nth frame we have a model M(N) which describes a configuration of these points in terms of its X, Y and Z coordinates. The X and Y coordinates are measured from their two-dimensional projection onto the image plane (we assume orthographic projection) while the Z (depth values) cooordinates have to be estimated. We define $L_{i,j}^N$ as the squared distance between points i and j, for the Nth frame, so that

$$L_{i,j}^N = (X_i^N - X_j^N)^2 + (Y_i^N - Y_j^N)^2 + (Z_i^N - Z_j^N)^2,$$

and the sum of $L_{i,j}^N$ over all (P) points, we define as $\mathcal{L}^{\mathcal{N}}$, that is,

$$\mathcal{L}^{\mathcal{N}} = \sum_{i=1}^{P} \sum_{j=1}^{P} L_{i,j}^N.$$

Now we go to the next frame and calculate the same quantity $L_{i,j}^{N+1}$, which is also a function of the unknown (new) depth values $\{Z_i^N\}$. The difference between $\mathcal{L}^{N+1}$ and $\mathcal{L}^N$ is a measure in the change of the rigidity of the internal model. To calculate the new depth values $\{Z_{i,j}^N\}$ we minimize a cost function $\mathcal{C}^{N+1,N}$, defined as the square of the difference between $\mathcal{L}^{N+1}$ and $\mathcal{L}^N$. Thus, we minimize

$$\mathcal{C}^{N+1,N} = [\mathcal{L}^{N+1} - \mathcal{L}^N]^2$$

with respect to $\{Z_{i,j}^{N+1}\}$. Note that the $\{Z_{i,j}^N\}$ are known (they have been obtained by an analogous process for the Nth frame). Having calculated the new depth values, we move to the $(N+2)$th frame and do the same computation for $\{Z_{i,j}^{N+2}\}$, and so on until we obtain the correct depth values (which correspond to a global minimum of the cost function). The minimization process gives us a third degree polynomial equation in the $Z_{i,j}^N$'s.

At each step of the (internal) iteration proceedure, the depth values $\{Z_{i,j}^N\}$ are initial inputs for the next step. So, as the number of iterations increases the correct values for the depth of the P points are approached. In this process $\mathcal{C}^{N+1,N}$ is, in general, different from zero which means that during the updating procedure, the change of structure (according by the internal model) is non-rigid.

To allow semi-rigid motion (as described at the end of the previous sec-

tion) we modify the IRS to the ISRS (Incremental Semi-Rigidity Scheme). The IRS minimizes the cost function $\mathcal{C}^{N+1,N}$, given by equation (10), which basically measures the change in the (global) rigidity of the internal model, between the two frames. However, if we restrict the summation in the expression of $\mathcal{L}^N$ given by equation (9) to be only between points which are vertices of the same triangles, then the difference between $\mathcal{L}^{N+1}$ and $\mathcal{L}^N$ will depend only on the sum over these vertices and we obtain the ISRS. We can rewrite equation (9) as

$$\mathcal{L}^N = \sum_{i,j} L_{i,j}^N,$$

where the summation is only taken over vertices i, j of the same triangles. So that the new $\mathcal{C}^{N+1,N}$ defined through (11) will now be a measure of the semi-rigidity of the structure.

The ISRS updates its internal model by minimizing (11) with respect to $\{Z_i^N\}$ thus enforcing only a local rigidity. In this sense, the ISRS has more flexibility to deal with object non-rigidity than does the IRS.

We should be carefull not to confuse the non-rigidity of the (three-dimensional) object with that of the internal model. Even if the object is rigid, the internal model will change in a non-rigid manner until it converges to the structure of the real object.

Specific computations

In the last two section we described the Regge calculus and the ISRS. Now we want to describe some specific computations that we did with different configurations built up of triangles.

Our general idea is to work with an arbitrarily triangulated surface in motion, allowing for bending deformations to take place. Initially we tested the ISRS for a rotating triangle. Afterwards we applied the same method to two adjacent triangles rotating and flexing over their common edge. Finally, we considered six triangles forming a closed sequence of adjacent elements which have three kinds of motion: (1) global rotation (each triangle rotates by the same amount), (2) an umbrella type of motion (where the six points on the perimeter have a oscilatory semirigid motion representing the closing or opening of a umbrella) and (3) motion with one edge on the perimeter changing its length by an oscilatory movement, thereby changing the curvature of the object.

For the single triangle rotating rigidly we obtained results similar to those of previous studies (Ullman 1984, Grzywacz and Hildreth 1985). We did not, however, observe any optimal angle of rotation under which the system best recovers structure.

We then proceeded to the two-triangle case. We gave this figure a combi-

nation of two kinds of motion, a global rotation around a fixed axis followed by a bending deformation over the common edge. The X (horizontal) and Y (vertical) cartesian coordinates parameterize the image plane, while the Z coordinate represents the (three-dimensional) depth value. We experimented with varying the axis of (global) rotation, starting with it pointing along the Y axis and then rotating it towards the Z axis by increments of 30 degrees. We also varied the amount of bending and rotation between times frames. The rotation varied from 10 to 60 degrees and the bending from 5 to 30 degrees.

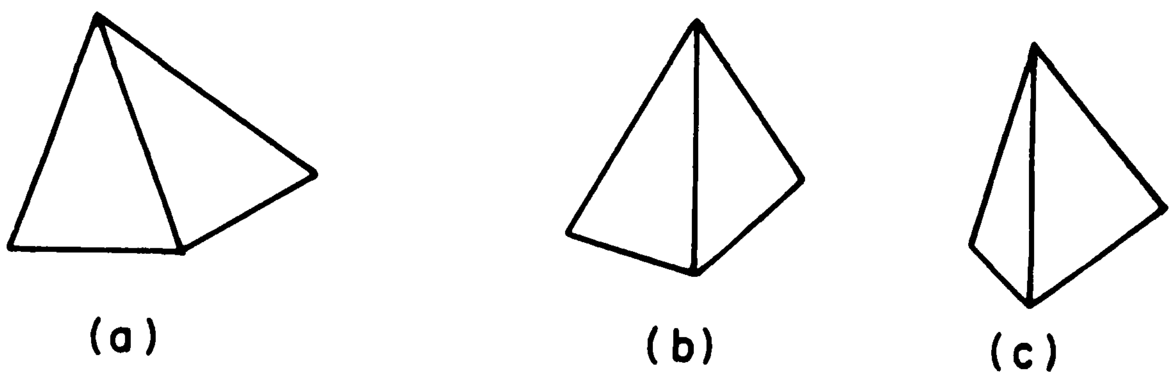


Figure 4 *Views of the two-triangle with different bending angles*

When the global rotation axis is at an angle of 30 or 45 degrees from the image plane (which means 60 or 45 degrees, respectively, to the Z axis) the results of the ISRS, the computed depth values, agree very well with the real values of

the simulation. On the other hand, for angles of 60 degrees (30 degrees with the Z axis), and especially 90 degrees (parallel to the Z axis), the recovery of structure from motion is not good. In the specific case of 90 degrees there is almost no information gained, since only part of the figure is visible in the sequence of frames. We also observed two well defined limits for the angle of (global) rotation in each time frame. It has a lower bound of about 5 degrees and an upper bound of about 90 degrees. We obtained the best results in the range between 30 and 60 degrees. It seemed that if these angles were too small or too large not enough information was available for the algorithm to use. Intuitively, if the global rotation is too small, recovery of the structure is unstable as too little new information is added between time frames. If the rotation is too large the new information is too different from the old one (you see the figure from a totally new point of view), so that the ISRS cannot build up a uniform model of the figure. This general properties of the IRS (and ISRS) were discussed in detail by Grzywacz and Hildreth (1985), where it is argued that if the size of the incremental rotation angles decreases then the deterioration of information is inversely proportional to the number of frames.

The best values of the bending angle (between each time frame) lay in the range of 10 to 30 degrees. Thus, since the bendings are made after the global

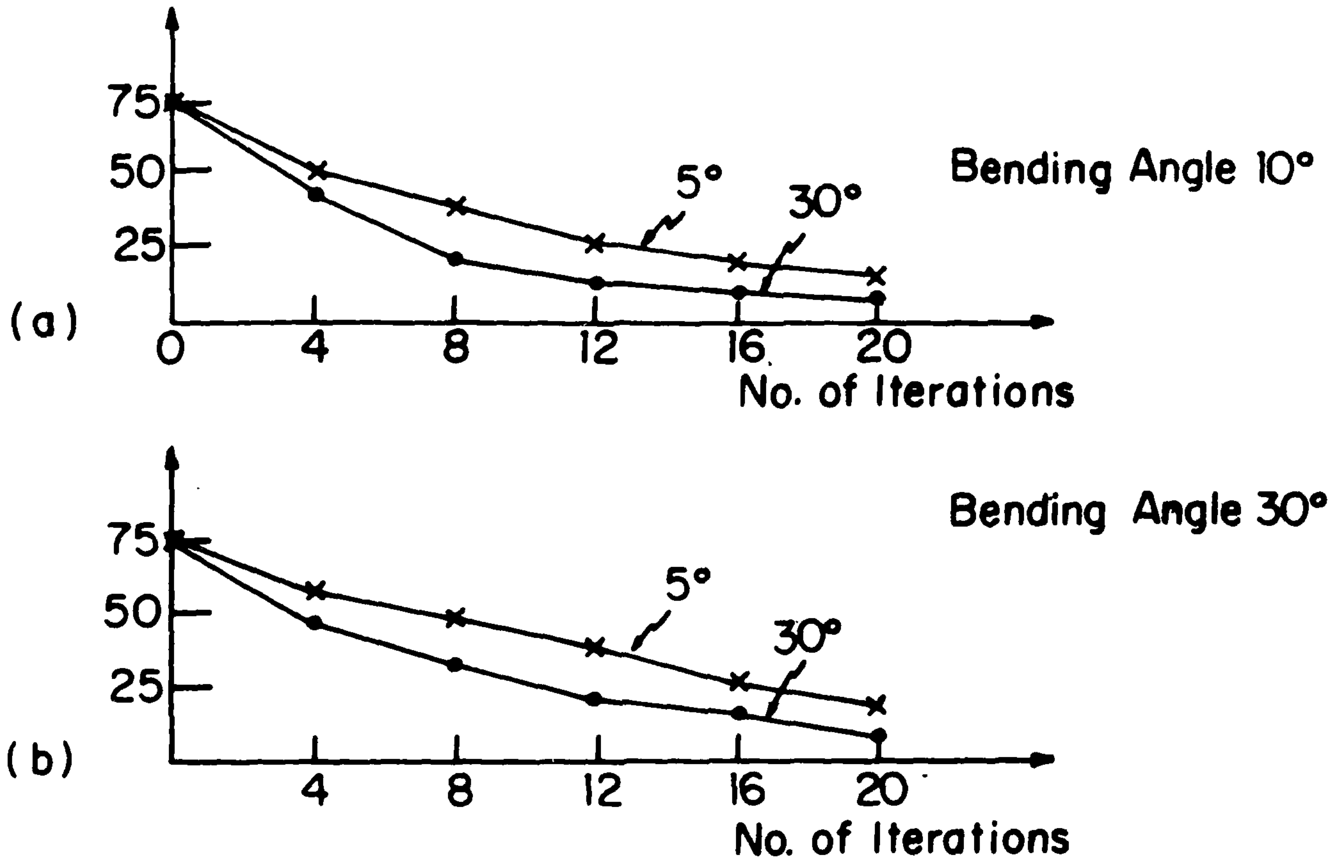


Figure 5 *Error graph for the two-triangles. The vertical axis shows the error function measure of the difference between the model and the stimulus. (a) and (b) have bending angles* 10 *and* 30 *degrees respectively. The rotation in each time frame is either* 5 *or* 30 *degrees.*

rotation of the figure, too large a bending angle can affect the robustness of the algorithm. So only limited amounts of flexing are allowed between time frames. Typically we did only one bending between each (global) rotation.

Next we considered six triangles performing a global rigid rotation around a fixed axis. Basically, the six-triangle figure consists of six adjacent triangles which all share a common vertex and any two adjacent triangles have always a common edge. The (global) rotation of this six-triangle is done with respect to a axis passing through the common vertex, and maintains all triangles with in a fixed (rigid) structure. The results we obtained are identical, in nature, to the ones obtained for the two triangles. More precisely if the orientation of the axis of (global) rotation, with respect to the image plane, is between 30 and 60 degrees, and the rotation angle (between time frames) is between 30 and 60 degrees then the ISRS algorithm is very efficient in recovering structure from motion.

For the next set of simulations we consider the six-triangle figure to simulate an "umbrella" type of motion. The "umbrella" type of motion consists in having three of the perimeter points (choosen in alternation) perform a oscillatory motion rather like the spokes of an umbrella being opened and closed. The motion of the three remaining points is determined uniquely by requiring the triangles to be rigid. This motion is illustrated in figure 6.

The umbrella type of motion can be defined in the following way. Let each point on the perimeter of the umbrella be described by a vector whose origin lies at a fixed point (the common vertex we introduced before) in space. We

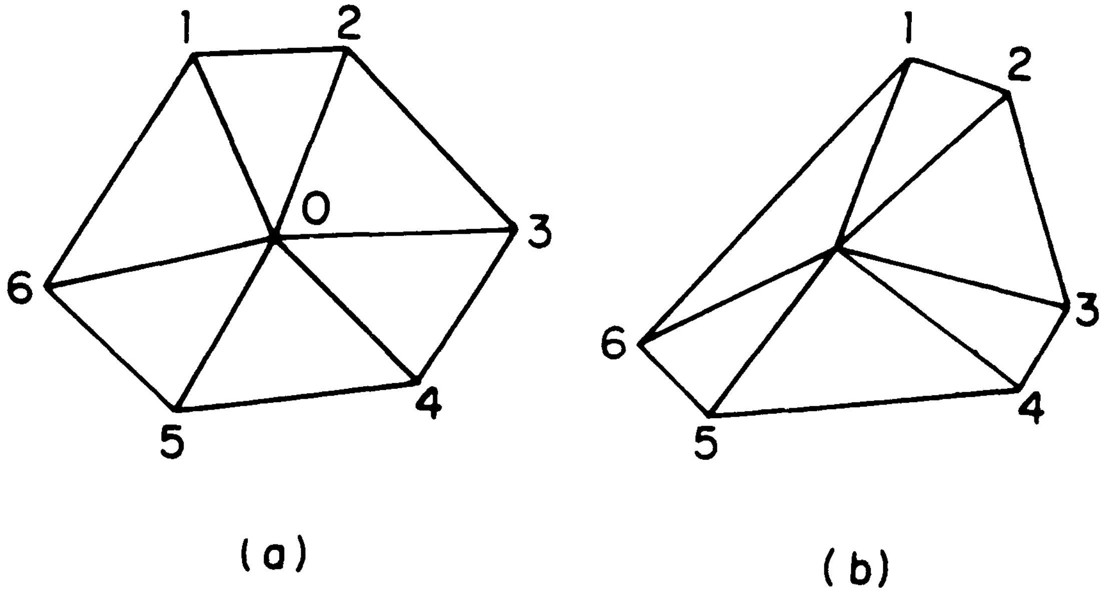


Figure 6 *"Umbrella" motion*

also define an axis with a specific orientation in the (three-dimensional) space (the handle of the umbrella). Of these six points only three can move freely in space so that the remaining points are constrained by the motion of the former ones. This can be easily understood by observing that the vectors of the constrained points can be expressed in terms of its adjacent unconstrained points by the following formula

$$R_i = \lambda R_{i-1} + \mu R_{i+1} + \nu R_{i-1} \wedge R_{i+1}$$

where R_i represents the unit vector defining the direction of the ith point (on the perimeter) in respect to the origin and "$\wedge$" is the cross product. Using simple algebraic manipulations, it can be shown that the parameters λ, μ and

ν are, respectively, given by

$$\lambda = \cos\alpha_{i-1,i} - R_{i-1} \cdot R_{i+1} \cos\alpha_{i,i+1} det$$

$$\mu = \cos\alpha_{i,i+1} - R_{i-1} \cdot R_{i+1} \cos\alpha_{i-1,i} det$$

and

$$\nu = \pm\sqrt{1 - \lambda^2 - \mu^2 - 2\lambda\mu R_{i-1} \cdot R_{i+1} det}$$

where

$$det \equiv 1 - (R_{i-1} \cdot R_{i+1})^2$$

The symbol "$\cdot$" represents the dot product and $\alpha_{i,i+1}$ represents the angle between vectors R_i and R_{i+1}. Notice that ν is determined up to a sign.

In this way, we label the six vertices of the perimeter of the umbrella from 1 to 6, and allow points 1, 3 and 5 to move independently (in a globally coherent way consistent with the triangles being rigid), with points 2, 4 and 6 satisfying the contraint (12). For the "umbrella" motion the three unconstrained (perimeter) points move by discrete changes of the polar angles ϕ_i (the angles between the vectors joining the points to the center and the handle of the umbrella). The value of these polar angles was constrained to lie between 60 and 120 degree. We varied these angles by different increments varying from 5 to 15 degrees.

For this stimuli it was more difficult to obtain a correct answer to the depth values. This showed itself in the difficulty of obtaining the global min-

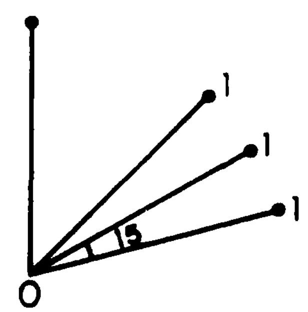


Figure 7 *Motion of one vertex*

ima of the cost function between frames. The system often got trapped in local minima during gradient descent. Some of these local minima had simple interpretations. For example some corresponded to a depth reversal of part of the structure (of course depth reversal of the whole structure is a possible ambiguity when orthographic projection is used). These particular minima could be removed by simple heuristics; one could find the endpoint of a minimization, flip the sign of a depth value and see if this reduced the energy. If it did, gradient descent could be restarted from this configuration. Not all minima, however, could be removed in this way. Even if the global minima was always found, which we could do by an interactive algorithm, the ISRS

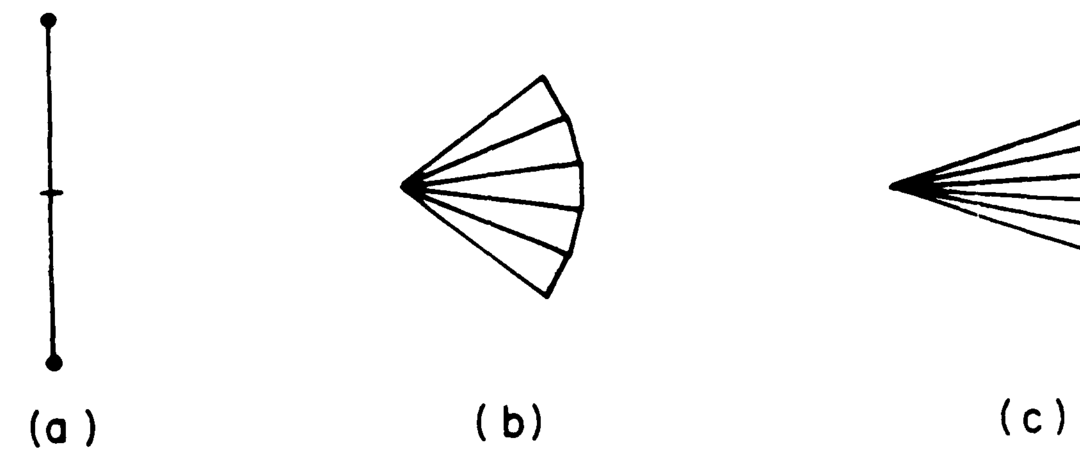


Figure 8 *Side views of the "umbrella"*

would not always converge to the correct result. Moreover, it would sometimes approach the right result and then diverge from it. Thus it seemed to display a number of the pathologies of the IRS (Grzywacz and Hildreth 1985).

We simulated the motion on the screen of a Symbolics LISP machine (with the help of V. Inada). Informal psychophysics suggested that humans also have difficulty estimating the depth for these stimuli, although they usually got the correct qualitative result. This suggested testing how good the ISRS results were qualitatively. To do this we also displayed the images and the models resulting from the IRS from different viewpoints. If the viewpoint was the same as for the simulation then naturally the projections of the im-

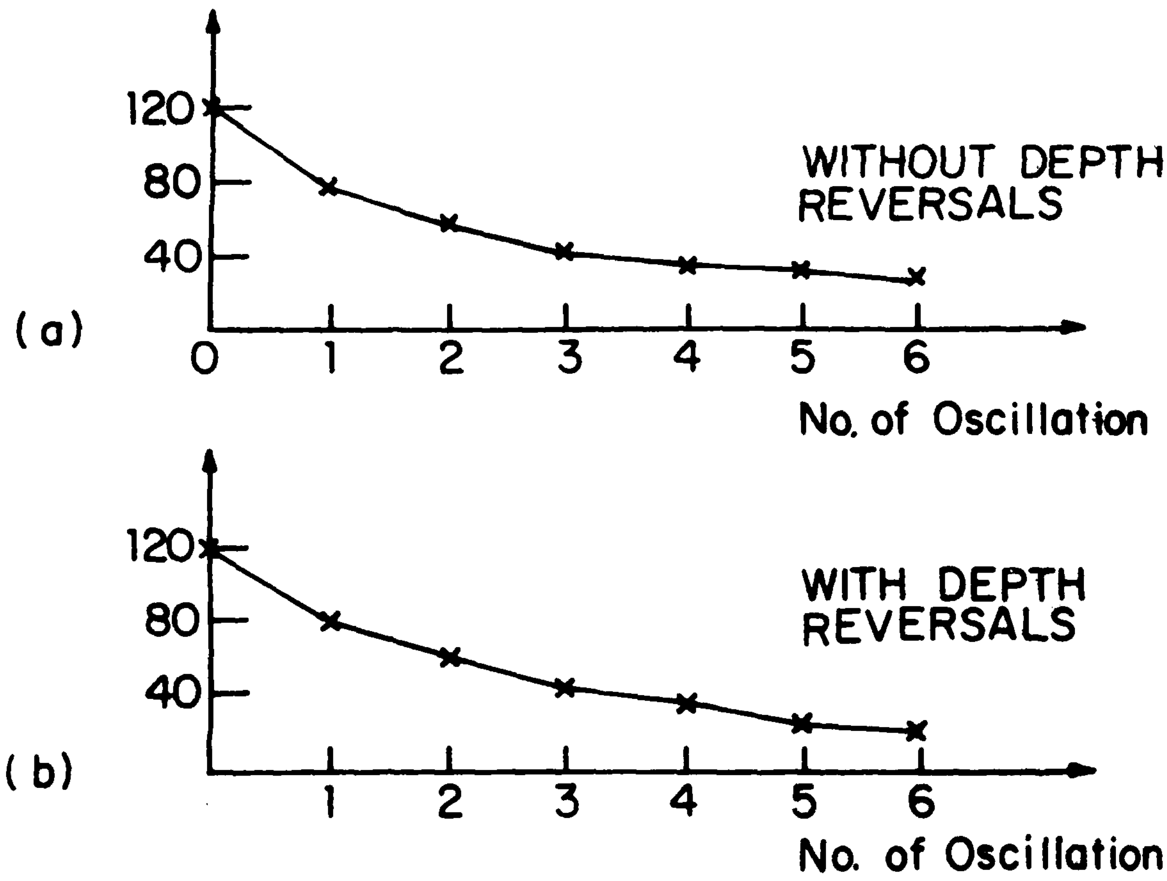


Figure 9 *The results for the umbrella. (a) and (b) show the convergence without and with depth reversals. The vertical axis shows the error function measure between the stimulus and the model.*

ages and the models was identical. By altering the viewpoint from this initial direction we could obtain a qualitative measure of the similarity between the image and the model. We found that, provided the axis of the umbrella is within about 30 degrees of the normal to the image plane and provided the result is viewed from a direction less than 30 degrees away from the normal to the image plane then the motion looks very similar to the simulated motion.

The "umbrella" motion is the most complicated one which has been simulated by either the IRS or the ISRS. We applied the IRS to some of the individual triangles of the umbrella and obtained poor results (worse than the estimates of the ISRS). These triangles move so that in some configurations their projected area is practically zero (see figure 6). We did not apply the IRS to the entire display, arguing that the large amount of rigidity of the entire structure would violate the assumptions of the IRS and prevent it from giving the correct answer. We know of no simulations of the IRS which work well under these conditions. This suggests that it is a better strategy to try an ISRS over the whole object than to apply the IRS to each part of the object seperately. The global effect of the ISRS creates a form of cooperation between parts of the object which might be badly estimated otherwise.

Finally we varied one edge of the perimeter of the six triangles with the others maintained constant. This caused a change of the Gaussian curvature

of the object. For small variations the results were good, but we did not do extensive experiments.

Conclusions

We showed that it is possible to recover structure from motion for figures constructed of triangles which are allowed to perform non-rigid motion with bending deformations. The basic idea is to segment the surface using Regge calculus, approximating it by a net of triangles with the curvature given by the sum of the deficit angles, and to recover structure from motion in terms of the Incremental Semirigidity Scheme. The vertices of the triangulation are used in the ISRS to obtain structure from motion.

The case of two triangles performing rigid global rotation followed by a local bending deformation shows that the ISRS algorithm is good when the axis of global rotation is close to the parallel position with respect to the image plane (parallel to the Y axis) and the angle of rotation lies in the range between 30 and 60 degrees. However, if this axis is close to being perpendicular to the image plane (parallel to the X axis), the algorithm shows poor performance. In addition, the bending angle has to be small (in the range between 10 and 30 degrees).

If we increase the number of triangles and the motion remains globally

rigid, then the recovery of structure from motion is best if the orientation of the rotation axis is close to parallel with respect to the image plane, as in the case of the two triangles. This means that increasing the number of points for this type of motion does not affect algorithmic robustness of the ISRS. On the other hand, if the six-triangle figures is allowed to perform the "umbrella" type of motion, then unless the position of the axis passing through the center of the common vertex is nearly parallel to the normal to the image plane, the algorithm does not perform well. Some informal psychophysics suggests that humans have some difficulty in correctly estimating the depth, but get the correct qualitative motion. The algorithm also often seemed to get the *correct qualitative motion.*

We have not yet discussed which points or elements on the surface are choosen to be the vertices of the triangulation scheme, nor how the correspondence between these points, for sucessive time frames, is done. Ullman (1984) suggested using features (detected by a suitable operator). For example, Hildreth (private communication) has used the texture features of a cup as input to the IRS. Detectable features of this type seem a natural choice for the vertices of the triangulation. It would be simple to adapt the IRS further to deal with objects made of rigidly moving subparts, for example rectangles, flexing at their joints. An extended model of this type might be able to explain Jo-

hansson's (1975) results for moving figures. In these experiments light sources were attached to the joints of moving figures and the correct (non-rigid) motion was retrieved. We should note that in this case the positions of the light sources suggest natural places to segment the object. The correspondence between sucessive time frames could be done by tracking, or by a minimal mapping scheme (1984). An interesting point is that the IRS depends only on the positions of the points and not on the lines connecting them. We did some informal psychophysics on the "umbrella" motion changing the triangulation by altering which points were connected by which lines (without changing the total number of points). These changes sometimes altered the depth perception of the object, in contrast to what the IRS would predict. These effects were only preliminary and need to be studied more systematically. However they suggest that the choice of triangulation is important.

We conclude from this that although the ISRS (and IRS) is good for a certain range of motions it is unable, at least without modifications, to cope with all motions. More studies are needed to check for which motions these schemes are effective. The IRS has mostly been demonstrated on constant rigid rotation about an axis and needs to be tested for a larger class of motions. We argue that the ISRS may be more effective than the IRS for non-rigid flexing motion since it has greater flexibility. Grzywacz and Hildreth have

suggested modifications to the basic IRS and report better results (private communication).

Acknowledgements

We would like to thank V. Inada, N. Grzywacz, E. Hildreth, T. Kanade and S. Ullman for interesting discussions. R.S.J. would like to thank CAPES(Brazilian Educational Agency) for financial support. A.L.Y. was employed by the Artificial Intelligence laboratory at MIT while most of this work was done. Support for the laboratory's artificial intelligence research is in part provided by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-85-K-0124.

References

A. D. Aleksandrov, and V. A. Zalgaller, **Intrinsic Geometry of Surfaces**, American Mathemathical Society, 15, 1967.

O. J. Braddick, **Low-level and high-level process in apparent motion**, Philos. Trans. R. Soc. London Ser. B, 290, 1980, 137-151.

J. L. Clathworthy, and J. P. Frisby, **Real and apparent movement: evidence for unitary mechanism**, Perception, 2, 1973, 161-164.

N. M. Grzywacz, and E. C. Hildreth, **The Incremental Rigidity Scheme**

for Recovering Structure from Motion: Positions vs. Velocity based Formulations, Artificial Intelligence Laboratory and Center for Biological Information Processing MIT, 1985, A.I. Memo 845, October 1985.

E. C. Hildreth, **The Measurement of Visual Motion**, MIT Press, Cambridge, Massachusetts, 1984.

E. C. Hildreth, and C. Koch, **The Analysis of Visual Motion: From Computational Theory to Neuronal Mechanisms**, Artificial Intelligence Laboratory and Center for Biological Information Processing, MIT, A. I. Memo 919, December 1986.

G. Johansson, **Visual Motion Perception**, Scientific American, 232, 6, 1975, 76-88.

J.J. Koenderink, and A. J. van Doorn, **Depth and Shape from Differential Perspective in the Presence of Bending Deformations**, J. Opt. Soc. Am. A, 3, 2, 1986.

D. Marr, **Vision A Computational Investigation into the Human Representation and Processing of Visual Information**, W. E. Freeman and Co., San Francisco,1982.

D. Marr, and E. C. Hildreth, **Proceedings of the Royal Society of London B**, 207, 1980, 187-217.

T. Poggio, and V. Torre, **Ill-Posed Problems and Regularization**

Analysis in Early Vision, Artificial Intelligence Laboratory and Center for Biological Information Processing, MIT. A.I. Memo 773, April 1984.

T. Regge, **General Relativity without Coordinates**, Il Nuovo Cimento, 19, 3, 1961, 558-571.

S. Ullman, **The Interpretation of Visual Motion**, MIT Press, Cambridge, Massachusetts, 1979.

S. Ullman, **Maximizing Rigidity:the Incremental Recovery of 3-D Structure from Rigid and Rubbery motion**, Perception, 13, 1984, 255-274.

H. Wallach, and D. N. O'Connell, **The Kinetik Depth Effect**, Journal of Experimental Psychology, 45, 1953, 4-205.

M. Wertheimer, **Experimentelle Studien ueber das Sehen von Bewegungen**, Zeitschrift für Psychologie, 61, 1912, [illegible]

END
DATE
FILMED
DTIC
4/88